EXHIBIT G

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 276, No. 20, Issue of May 18, pp. 16919–16930, 2001
Printed in U.S.A.

Localization and Stability of Introns Spliced from the *Pem* Homeobox Gene*

Received for publication, June 13, 2000, and in revised form, February 20, 2001
Published, JBC Papers in Press, February 23, 2001, DOI 10.1074/jbc.M005104200

Jade Q. Clement, Sourindra Maiti, and Miles F. Wilkinson‡

From the Department of Immunology, the University of Texas, M. D. Anderson Cancer Center, Houston, Texas 77030

RNA splicing generates two products in equal molar amounts, mature mRNAs and spliced introns. Although the mechanism of RNA splicing and the fate of the spliced mRNA products have been well studied, very little is known about the fate and stability of most spliced introns. Research in this area has been hindered by the widely held view that most vertebrate introns are too unstable to be detectable. Here, we report that we are able to detect all three spliced introns from the coding region of the *Pem* homeobox gene. By using a tetracycline (tet)-regulated promoter, we found that the half-lives of these *Pem* introns ranged from 9 to 29 min, comparable with those of short lived mRNAs such as those encoding c-*fos* and c-*myc*. The half-lives of the *Pem* introns correlated with both their length and 5′ to 3′ orientation in the *Pem* gene. Subcellular fractionation analysis revealed that spliced *Pem* introns and pre-mRNA accumulated in the nuclear matrix, high salt-soluble, and DNase-sensitive fractions within the nucleus. Surprisingly, we found that all three of the spliced *Pem* introns were also in the cytoplasmic fraction, whereas *Pem* pre-mRNAs, U6 small nuclear RNA, and a spliced intron from another gene were virtually excluded from this fraction. This indicates either that spliced *Pem* introns are uniquely exported to the cytoplasm for degradation or they reside in a unique soluble nuclear fraction. Our study has implications for understanding the regulation of RNA metabolism, as the stability of introns and the location of their degradation may dictate the following: (i) the stability of nearby mRNAs that compete with spliced introns for rate-limiting nucleases, (ii) the rate at which free nucleotides are available for further rounds of transcription, and (iii) the rate at which splicing factors are recycled.

Although spliced introns and mRNAs are spliced from pre-mRNA in equal molar amounts, far less is known about the fate and stability of introns than of mRNAs (1). This may be because most introns have no known function. Apart from a few "specialty introns" that mediate antisense regulation (2), enhance transcription (3), encode small nucleolar RNAs that play a role in rRNA processing (4, 5), or encode proteins that mediate RNA splicing and transposition (6), it is widely assumed that most introns have no specific functional attributes. Instead, evidence suggests that most introns in modern organisms arose serendipitously as a result of the following two processes: the shuffling of small intronless primordial genes to generate large genes with introns (the intron-early theory), and the introduction of introns into intronless genes by transposition-type events (the intron-late theory) (7).

What are the consequences of the fact that modern eukaryotic genomes are saddled with large numbers of introns? First, evidence suggests that introns play a major evolutionary role in shaping protein function by virtue of their ability to promote exon shuffling (8). Second, it is clear that introns are necessary for the efficient expression of most mammalian genes (9–12). Presumably, the presence of introns in pre-mRNA sends the newly transcribed RNA down the appropriate pathway to permit high level accumulation of spliced mRNA.

What is less clear is the importance of introns *after* they are spliced out of pre-mRNA. Intron sequences exceed the length of exon sequences in most vertebrate genes (13, 14), and thus the major component of spliced transcription units that are ultimately degraded is introns. The rate of this degradation may influence the levels of nucleotides available for further rounds of transcription. In addition, intron turnover rate may influence RNA splicing. U2, U5, and U6 small nuclear ribonuclear proteins and Ser-Arg-containing proteins remain bound to released intron lariats (15), and therefore the metabolism of introns may have a significant influence on the recycling of these splicing factors. Vertebrate cells execute complex constitutive and alternative splicing events that depend on precise concentrations of splicing-regulatory factors. The rate of intron turnover may significantly affect the availability of these splicing factors, thereby regulating many cellular processes in multicellular organisms.

The importance of efficient intron turnover is supported by studies on organisms deficient in debranchase, the enzyme that specifically cleaves the 2′–5′-phosphodiester bond at the branch site of intron lariats (16). *Schizosaccharomyces pombe* deficient in debranchase because of a null mutation accumulate high levels of intron lariats and exhibit a severe growth defect (17). That this growth defect in fission yeast may stem from a toxic build up of undegraded introns is suggested by a comparison with the budding yeast *Saccharomyces cerevisiae*, which contains ~40 times fewer introns than *S. pombe* and does not display obvious phenotypic defects when rendered debranchase-deficient (18). Vertebrate genomes contain even more introns than *S. pombe*, and therefore an interference with intron decay is also likely to have serious consequences in vertebrates.

Given the probable importance of intron turnover, it is surprising that very little is known about this topic. In part, this deficiency may reflect the fact that very few spliced nuclear pre-mRNA introns have been detected in vertebrate cells. Most of the spliced introns that have been observed *in vivo* are

* This work was supported by National Institutes of Health Grant CA78023 and National Science Foundation Grant MCB-9307963. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Immunology, the University of Texas M. D. Anderson Cancer Center, Box 180, 1515 Holcombe Blvd., Houston, TX 77030. Tel.: 713-794-5526; Fax: 713-745-0846; E-mail: mwilkins@mail.mdanderson.org.

derived from strongly transcribed genes. For example, excised introns from the adenovirus-2 E2A pre-mRNA were identified in HeLa cells treated with cycloheximide to increase the rate of transcription (19). A β-globin intron was detected from rabbit liver, a source rich in β-globin pre-mRNA (20). A spliced immunoglobulin-κ intron was observed in a stimulated plasmacytoma cell line that transcribes high levels of an immunoglobulin-κ gene (21).

A widely held belief is that spliced introns accumulate at low levels because they are rapidly degraded (within seconds) at their site of origin in the nucleus (14, 22). However, there is little direct evidence to support this view. The only mammalian intron whose fate and stability have been examined in detail is the $IVS1_{C\beta1}$ intron from a mouse T-cell receptor (TCR)[1]-β gene. This intron is easily detectable by the relatively insensitive Northern blot procedure, despite being generated from only a modestly transcribed gene (23). The half-life of $IVS1_{C\beta1}$ in HeLa cells was determined to be 6 min (24), which is much longer than was originally proposed for introns in general (14, 22). Most spliced $IVS1_{C\beta1}$ lariats were found to be in the nuclear compartment, consistent with their origin and degradation in the nucleus (24).

In the present communication, we report on the fate, localization, and stability of spliced introns from the *Pem* homeodomain gene, a mammalian gene that contains typical mammalian introns ranging in length from ~0.2 to 2.4 kb (25, 26). Unlike the other genes from which spliced introns have been studied, the *Pem* gene is not expressed in a cell type-specific manner. Instead, *Pem* is expressed by different cell types in many different fetal and adult tissues, as well as in tumor cells from several different lineages (25–32). Because the *Pem* gene is not cell type-specific and its introns appear to be typical, we hypothesized that *Pem* would be a good candidate to provide information on the metabolism of mammalian introns in general.

We used Northern blot analysis and RPA to detect spliced *Pem* introns, because these approaches have several advantages, including their ability to provide a quantitative measure of RNA levels and their ability to resolve (by size) spliced introns from intron-containing pre-mRNAs. Other methods used to identify, localize, or measure the stability of intron-bearing RNAs (*e.g.* pulse-chase analysis, reverse transcription-polymerase chain reaction (PCR) analysis, and *in situ* hybridization) are either not quantitative or do not distinguish between spliced introns and pre-mRNAs (33–36). We used several independent approaches to verify that we had identified *bona fide* spliced *Pem* introns. Our analysis of the three introns in the *Pem* coding region revealed that they had a range of half-lives that were even longer than that of the only other previously analyzed vertebrate intron, $IVS1_{C\beta1}$. The spliced *Pem* introns were found in several fractions of the nuclear compartment, consistent with their origin in the nucleus. Surprisingly, the *Pem* introns were also present in the cytoplasmic fraction of several different cell types. Together, these findings may have important implications for RNA metabolism in general.

EXPERIMENTAL PROCEDURES

Cells—The cells used for our study were the kind gift of the following investigators: the human HeLa cell line (Susan Berget, Baylor School of Medicine), the rat PS-1 prostate mesenchymal cell line (David Rowley, Baylor School of Medicine), the rat FRTL-5 thyroid cell line (Christine Spitzweg, Mayo Clinic), and the rat 208-F, Rat-1/S (*src*-transformed), and Rat-1/R (*ras*-transformed) fibroblast cell lines (Bruce Magun, Oregon Health Sciences University). The rat McA-RH8994 liver cell line was obtained from the American Type Culture Collection. The myelin basic protein-reactive rat T-cell line was obtained from Arthur Vandenbark and Halina Offner (Oregon Health Sciences University).

RNA Isolation and Subcellular Fractionation—Nuclear, cytoplasmic, and total cellular RNAs were isolated as described previously (30, 37). For each cell line, the length of time in the lysis buffer (0.6% Nonidet P-40, 0.15 M NaCl, 10 mM Tris (pH 8), 0.1 mM EDTA) was optimized (between 3 and 10 min) such that >90% of cells exhibited disrupted cell membranes as judged by staining with toluidine blue O. Under the conditions chosen, we found that none of the cell lines displayed nuclear lysis, as judged by microscopic visualization of the released nuclei and by the fact that there was no evidence of viscous nuclear pellets (disrupted nuclei leak DNA, which causes aggregation of nuclei and increased viscosity). Further assurance of purity was demonstrated by methylene blue staining of nylon membranes onto which RNA from the nuclear and cytoplasmic fractions had been transferred (38). We found that cytoplasmic RNA had only mature 18 S and 28 S rRNA, whereas nuclear RNA also had 32 S and 45 S precursor rRNAs, in accordance with its expected properties (38). In cases in which a nuclear wash was performed, the nuclei were resuspended in 0.5% sodium deoxycholate in lysis buffer and immediately centrifuged.

Subcellular fractionation of HeLa cells was performed as described (39). In brief, trypsinized cells were fractionated into cytoplasmic (after Nonidet P-40 lysis), nuclear membrane (after Nonidet P-40 and sodium deoxycholate wash), chromatin-associated (after DNase I digestion), high salt-soluble (after incubation with a high concentration of NaCl), and nuclear matrix (remaining pellet) fractions. The RNA from these five fractions was purified by using method 1 described in Ref. 37.

Northern Blot and RNase Protection Analysis (RPA)—For Northern blot analysis, RNA samples were electrophoresed in agarose or polyacrylamide gels, blotted, and hybridized as described (24). Before prehybridization, all blots were stained with methylene blue to demonstrate RNA integrity, show equivalent loading and transfer, and judge the purity of nuclear and cytoplasmic RNA (38). Standard RNA molecular weight markers (RNA Molecular Weight marker I (Roche Molecular Biochemicals)) or the 0.16–1.77- and 0.24–9.5-kb RNA ladders (Life Technologies, Inc.) were used to determine the size of RNA transcripts. The DNA probes used for hybridizing the Northern blots were generated by PCR using the DNA oligonucleotides listed in Table I. Table II provides a list of the DNA probes. The names of the probes correspond to the exons (E) or introns (I) in the probes. Most DNA probes were labeled using [^{32}P]dATP and a Roche Molecular Biochemicals random-primed DNA labeling kit. To increase the specific activity of the IVS1 probe, both [^{32}P]dATP and [^{32}P]dCTP were used. Band intensities of Northern blots were determined by phosphorimage analysis or densitometry.

We determined the half-lives of the *Pem* introns by the same procedures we described previously (24). In brief, when the cells reached 70–80% confluence, they were incubated with 1 μg/ml tet (Sigma) for the times indicated in the figures. Total cellular, nuclear, or cytoplasmic RNAs were isolated at various time intervals after the addition of tet. The kinetic loss of RNA was monitored by Northern blot analysis. Levels of *Pem* RNAs for each time point were normalized against the corresponding levels of cyclophilin mRNA. The RNA half-lives were assessed by least squares linear regression analysis as described previously (24).

RPA was performed essentially as we have done before (25, 30) using riboprobes synthesized with T3 RNA polymerase and labeled with [^{32}P]UTP. The Pem IVS2/exon probe, which contains 234 nt of IVS2 and 106 nt of the upstream exon, was made by PCR using the oligonucleotides N and O, the latter includes a T3 polymerase-binding and initiation site. In an analogous fashion, the IVS3/exon probe, which contains 333 nt of IVS3 and 46 nt of the upstream exon, was generated using the primers P and Q. The size of protected bands was determined by comparison with a [^{32}P]UTP-labeled RNA Century ladder composed 100-, 200-, 300-, 400-, and 500-nt RNA transcripts (Ambion, Inc., Austin, TX). Band intensities of RPAs were measured using an Instant Imager (Packard Instrument Company, Meriden, CT), which directly measures radioactivity.

Primer Extension Analysis and Debranching of Lariat Introns—Primer extension experiments were performed as described previously (24). Oligonucleotide J (Table I), which is complementary to the 3′ end of the IVS2, was annealed to total cellular RNA from HeLa cells stably transfected with pTARP. The sequencing ladder was generated by the standard dideoxy-mediated chain termination DNA sequencing method (40). Oligonucleotide M (Table I), which is complementary to the 5′ end

[1] The abbreviations used are: TCR, T-cell receptor; IVS, intervening sequence; PCR, polymerase chain reaction; RPA, ribonuclease protection analysis; snRNA, small nuclear RNA; tet, tetracycline; nt, nucleotide; kb, kilobase pair.